S.I. Musaeva [1]
R.I. Nugmanov [1]
A. I. Lin [1,2]
T.I. Madzhidov [1]
O. Klimchuk [2]
I. S. Antipin [1]
A. Varnek [2]

ELECTRONIC GREEN'S REACTIVITY CHART BASED ON BIG DATA ANALYSIS

[1] A.M. Butlerov Chemical Institute, Kazan Federal University, Kremlevskaya 18, Kazan, 420008, Russia;
[2] University of Strasbourg, 4, Blaise Pascal str., Strasbourg, 67081, France;

stsouko@live.ru

Protective Group (PG) is a special chemical group which protects Functional Group (FG, OH group on the Figure below) during the synthesis. These groups help chemists to perform their reactions without any risk to transform undesirable parts of molecule.

The question is how to assess PG reactivity for effective synthesis planning if its reactivity depends on used conditions (catalyst, for example). Usually, the chemist relies either on his/her own experience or on the information from the literature, such as "Greene's Protective Groups in Organic Synthesis" [1]. Unfortunately, in the latter case, it is not clear how the authors assessed the PG reactivity in simple and selective reactions with molecules contained several PG.

In this presentation we demonstrate clear advantage of chemoinformatics approaches to assess Protective Groups reactivity over expert knowledge acquired from the "manual" analysis of the literature. Using new smart substructure search we classified transformations (PG is cleaved or remains) in simple and selective reactions taken from Reaxys database. To analyze reaction conditions we created the dictionary of synonyms contained "raw" and "standard" compounds names. Applying the Condensed Graph of Reaction (CGR) approach [2] we have performed a statistical analysis. Its results we compared with the Green's reactivity charts and have found some disagreements.

1. Wuts P.G.M. Greene's Protective Groups in Organic Synthesis, 2014, Wiley.
2. Hoonakker F. *Int. J. Artif. Intell. Tools*, 2011, **20**(2): 253–270.

The research was supported by Russian Scientific Foundation, grant 14-43-00024.